Journal of Computational and Applied Mathematics 232 (2009) 295–304

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Robust portfolio selection based on asymmetric measures of variability of stock returns

Wei Chen *, Shaohua Tan

State Key Laboratory of Machine Perception, Peking University, Beijing 100871, China

ARTICLE INFO

Article history:
Received 25 November 2008
Received in revised form 3 February 2009

Keywords:
Interval random uncertainty set
Interval random chance-constrained programming
Robust portfolio selection

ABSTRACT

This paper addresses a new uncertainty set—interval random uncertainty set for robust optimization. The form of interval random uncertainty set makes it suitable for capturing the downside and upside deviations of real-world data. These deviation measures capture distributional asymmetry and lead to better optimization results. We also apply our interval random chance-constrained programming to robust mean-variance portfolio selection under interval random uncertainty sets in the elements of mean vector and covariance matrix. Numerical experiments with real market data indicate that our approach results in better portfolio performance.

1. Introduction

In 1952, Markowitz [1] published his pioneering work that paved the foundation for modern portfolio analysis. Despite the theoretical success of Markowitz's model, the consideration of estimation risk and model risk has grown in importance. Since the estimates of the market parameters are reflected by statistical errors, the solutions of portfolio optimization are often very sensitive to perturbations in the market parameters. As in practice expected returns and asset covariances cannot be measured exactly but have to be estimated with large errors sometimes. It is important that we take uncertainty resulting from estimation errors into account.

Most recently, researchers have incorporated the uncertainty introduced by estimation errors directly into the portfolio optimization process by robust optimization introduced in [2,3]. In this case, the inputs are not classical ones, such as expected returns and covariances, but rather uncertainty sets (see, for example, [4–6]). In [7], Ben-Tal and Nemirovski summarized two frequently used uncertainty sets:

(1) "Unknown-but-bounded" uncertainty set. There are two forms here. First, uncertainty set defines an N-dimensional box [8]. It considers possible deviations of the N uncertain parameters from their expected values. For example,

$$U(\hat{u}) = \{u : |u_i - \hat{u}_i| \le \delta_i, i = 1, \dots, N\},$$

where u_i is the expected return for asset i, and $\hat{u}_i$ is the estimate of u_i. Second, uncertainty set defines an N-dimensional ellipsoid [8]. For example,

$$U(\hat{u}) = \{u : (u - \hat{u})' \Sigma_u^{-1} (u - \hat{u}) \le \delta^2\},$$

where Σ_u is the covariance matrix of estimation errors for the vector of expected returns u.

(2) "Random symmetric" uncertainty set. For example, the estimated values $\hat{u}_i$ are obtained from the mean values u_i by random perturbations:

$$\hat{u}_i = (1 + \epsilon_i) u_i,$$

where ϵ_i are independent random variables symmetrically distributed in the interval $[-1, 1]$.

* Corresponding address: State Key Laboratory of Machine Perception, Peking University, Room 2314, Beijing 100871, China. Tel.: +86 010 62756374.
E-mail addresses: pekingchenwei@hotmail.com, chenwei@cis.pku.edu.cn (W. Chen).

0377-0427/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2009.06.010

Though these two uncertainty sets are frequently used, they have two serious disadvantages. First, it is difficult to collect all information to determine the precise bounds of "unknown-but-bounded" uncertainty set in practice. Sometimes only the distributions of the bounds can be found from historical data. In this case, the "unknown-but-bounded" uncertainty set is actually fluctuant instead of stable. Therefore, the variability of bounds cannot be ignored. Second, the assumption of symmetric distribution is also limiting in many applications especially in financial modeling in which distributions are often known to be asymmetric.

In this paper, we introduce a novel uncertainty set: interval random uncertainty. In our early paper [9], we presented the definition of interval random variable and several interval random programming models. Roughly speaking, an interval random variable is an interval with random fluctuant bounds. For example,

$$\xi_i = [u_i - \theta_i^1, u_i + \theta_i^2],$$

where ξ_i are interval random variables. The mean values u_i are mean-point of ξ_i. Random variables θ_i^1 and θ_i^2 are downside and upside deviations of ξ_i around mean values respectively. Interval random variable ξ_i consider the variability of bounds and asymmetric measures of variability for the distribution of data simultaneously. Hence, it is a good idea to introduce interval random variable as uncertainty set. Then we apply our interval random chance-constrained programming to robust mean-variance portfolio selection. We also address the way to generate uncertainty set based on moving averages. Finally, A hybrid-intelligent algorithm is applied to solve the robust portfolio model. some computational results are discussed that demonstrate the potentially significant economic benefits of investing in portfolios computed using classical models and the model introduced here. The robustness is achieved at relatively high performance and low cost.

The rest of this paper is organized as follows. Section 2 presents basic definitions of interval random variable. In Section 3, interval random chance-constrained programming model is discussed. Section 4 proposes the robust portfolio selection model using interval random chance-constrained programming. Section 5 presents results of some computational experiments with our robust portfolio model. Finally, a few concluding remarks are given in Section 6.

2. Interval random uncertainty set

A detailed description of interval random variable has been made in our early paper [9]. Roughly speaking, an interval random variable is a measurable function from a probability space to a collection of closed intervals. In other words, an interval random variable is a random variable taking interval values. Let $\mathcal{I}$ be a collection of closed intervals. For our purpose, we use the following definition of interval random variable [9].

Definition 1. Let $(\Omega, \mathcal{A}, Pr)$ be a probability space. An interval random variable is a function $\xi : \Omega \to \mathcal{I}$ such that $\xi(\omega) = [\underline{\xi(\omega)}, \overline{\xi(\omega)}]$ is a measurable function of ω.

For example, let $(\Omega, \mathcal{A}, Pr)$ be a probability space. Let $\Omega = \omega_1, \omega_2, \dots, \omega_m$, and $\mu_1, \mu_2, \dots, \mu_m$ be closed intervals in $\mathcal{I}$. Then the function

$$\xi(\omega) = \begin{cases} \mu_1, & \text{if } \omega = \omega_1, \\ \mu_2, & \text{if } \omega = \omega_2, \\ \vdots & \\ \mu_m, & \text{if } \omega = \omega_m \end{cases}$$

is clearly an interval random variable.

Definition 2. Let $f : \mathcal{I}_n \to \mathcal{I}$ be a function over the n-dimensional Euclidean space and ξ_i be interval random variable defined on $(\Omega_i, \mathcal{A}_i, Pr_i)$, $i = 1, 2, \dots, n$, respectively. Then, $\xi = f(\xi_1, \xi_2, \dots, \xi_n)$ is an interval random variable on $(\Omega_1 \times \cdots \times \Omega_n, \mathcal{A}_1 \times \cdots \times \mathcal{A}_n, Pr_1 \times \cdots \times Pr_n)$, defined by $\xi(\omega_1, \omega_2, \dots, \omega_n) = f(\xi_1(\omega_1), \xi_2(\omega_2), \dots, \xi_n(\omega_n))$, for all $(\omega_1, \omega_2, \dots, \omega_n) \in \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$.

The assumption of symmetric distribution of traditional uncertainty sets makes no distinction between downside and upside deviations. But in real world, the distributions of returns of assets are often known to be asymmetric. The form of interval random variable makes it suitable for capturing the downside and upside deviations of asset returns. We assume that the uncertainty set for the expected return u_i of asset i and covariance δ_{ij} of asset i and asset j take the form of interval random variables:

$$U(u_i) = [m_i - \theta_i^1, m_i + \theta_i^2]$$
$$U(\delta_{ij}) = [m_{ij} - \tau_{ij}^1, m_{ij} + \tau_{ij}^2]$$

where m_i are the mean values of returns. Random variables θ_i^1 and θ_i^2 are downside and upside deviations for the distribution of returns respectively. m_{ij} are the mean values of covariances. Random variables τ_{ij}^1 and τ_{ij}^2 are downside and upside deviations for the distribution of covariances respectively. Hence, they enable us to capture the asymmetry of asset returns in order to obtain better solutions that satisfy chance constraints. If $\theta_i^1 = \theta_i^2$ or $\tau_{ij}^1 = \tau_{ij}^2$, then the interval random uncertainty sets become symmetric.

Let x_i denote the proportion of the portfolio to be invested in asset i. The expected return of portfolio is defined as $R = \sum_{i=1}^{i=N} x_i u_i$. And the variance of portfolio is $V = \sum_{i=1}^{i=N} \sum_{j=1}^{j=N} x_i x_i \delta_{ij}$. The mean-variance optimization problem can be expressed as follows:

$$\begin{aligned} &\max_{\{u_i \in U(u_i), \delta_{ij} \in U(\delta_{ij})\}} && \sum_{i=1}^{N} u_i x_i - \lambda \sum_{i=1}^{i=N} \sum_{j=1}^{j=N} x_i x_i \delta_{ij} \\ &\text{s.t.} && \sum_{i=1}^{N} x_i = 1 \\ &&& x_i \geq 0. \end{aligned} \tag{1}$$

For each given decision vector x, it is meaningless to maximize the objective function before we know the values of u_i and δ_{ij}. Also, we cannot judge whether or not a decision vector x is feasible before we know the precise values of u_i and δ_{ij}. To solve this problem, we will present mathematically meaningful interval random chance-constraint programming models in the next section.

3. Interval random programming

3.1. Chance of interval random event

First, let us recall the definition of function $\mathcal{F}ea(A \leq B)$ [9] which represents the grade of feasibility of the interval event 'A to be less than or equal to B'.

Definition 3. Let I be the set of all closed intervals on the real line $\mathfrak{R}$. Let $A \in I, B \in I, A = [\underline{a}, \overline{a}], B = [\underline{b}, \overline{b}]$. We define a feasibility function $\mathcal{F}ea: I \times I \to [0, 1]$ such that

$$\mathcal{F}ea(A \leq B) = \begin{cases} 1, & \overline{a} \leq \underline{b} \\ 1, & m(A) = m(B) \\ \dfrac{m(B) - m(A)}{w(B) + w(A)}, & m(A) < m(B) \text{ and } \overline{a} > \underline{b} \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where $m(A)$ and $w(A)$ are the mid-point and half-width of interval A, $m(A) = \frac{\underline{a}+\overline{a}}{2}$, $w(A) = \frac{\overline{a}-\underline{a}}{2}$. Similarly, $m(B) = \frac{\underline{b}+\overline{b}}{2}$, $w(B) = \frac{\overline{b}-\underline{b}}{2}$. At least one of A and B must be interval. If one of them is a real number, for example B, then $m(B) = \underline{b} = \overline{b}$, $w(B) = 0$.

If interval event ζ is composed of several interval events $(\zeta_1, \zeta_2, \ldots, \zeta_n)$, ζ is called a joint interval event. The feasibility function of joint interval event is defined as $\mathcal{F}ea(\zeta) = \inf_{i=1,\ldots,n} \mathcal{F}ea(\zeta_i)$.

Now, let us consider the chance of an interval random event. Like the chance of fuzzy random event [10], the chance for an interval random event is also a function. Generally, we have the following definition of chance of interval random event [9].

Definition 4. Let $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ be an n-dimensional interval random vector and $f : \mathcal{I}^n \to \mathcal{I}$ be functions. Then the chance of interval random event characterized by $f(\xi) \leq 0$, is a function $\mathcal{C}h$ from $(0, 1]$ to $[0, 1]$ such that for any $\alpha \in (0, 1]$. We have

$$\mathcal{C}h\{f(\xi) \leq 0\}(\alpha) = \sup\{\beta | Pr\{\omega \in \Omega | \mathcal{F}ea\{f(\xi(\omega)) \leq 0\} \geq \beta\} \geq \alpha\}.$$

The chance $\mathcal{C}h\{f(\xi) \leq 0\}(\alpha)$ represents "the interval random event holds with feasibility $\mathcal{C}h\{f(\xi) \leq 0\}(\alpha)$ at the probability α".

3.2. Interval random chance-constrained programming

First, we consider the following generalized programming model with interval random variable:

$$\begin{cases} \max & f(x, \xi) \\ \text{s.t.} & g_j(x, \xi) \leq 0, \quad j = 1, 2, \ldots, m \\ x \in D \end{cases} \tag{3}$$

where x is the decision vector, and ξ is the interval random vector.

According to Definition 2, the return of $f(x, \xi)$ is an interval random variable. To measure the return of function $f(x, \xi)$, in [9] we define two critical values: optimistic value and pessimistic value.

Definition 5. For any given decision x

$$f_{\max} = \max\{r | \mathcal{C}h\{f(x, \xi) \geq r\}(\gamma) \geq \delta\}$$

is called the (γ, δ)-optimistic value to the return function $f(x, \xi)$, where $\gamma \in (0, 1]$, $\delta \in [0, 1]$.

Definition 6. For any given decision x

$$f_{\min} = \min\{r|\mathcal{C}h\{f(x,\xi) \le r\}(\gamma) \ge \delta\}$$

is called the (γ, δ)-pessimistic value to the return function $f(x, \xi)$, where $\gamma \in (0, 1]$, $\delta \in [0, 1]$.

It is naturally desired that the interval random constraints in (3) hold with feasibility β at probability α, where α and β are specified confidence levels, $\alpha \in (0, 1]$, $\beta \in [0, 1]$. Then, we have a chance constraint as follows:

$$\mathcal{C}h\{g_j(x,\xi) \le 0, j = 1, 2, \ldots, m\}(\alpha) \ge \beta. \tag{4}$$

This type of chance constraint is called a joint chance constraint. Sometimes, the following separate chance constraint is employed:

$$\mathcal{C}h\{g_j(x,\xi) \le 0\}(\alpha_j) \ge \beta_j, \quad j = 1, 2, \ldots, m$$

where α_j and β_j are confidence levels, $\alpha_j \in (0, 1]$, $\beta_j \in [0, 1]$ for $j = 1, 2, \ldots, m$.

If we want to maximize the optimistic value to the interval random return function subject to some chance constraints, then we have the following interval random maximax chance-constrained programming model:

$$\begin{cases} \max\limits_{x} \quad \max\limits_{\bar{f}} \quad \bar{f} \\ \text{s.t.} \qquad\qquad \mathcal{C}h\{f(x,\xi) \ge \bar{f}\}(\gamma) \ge \delta \\ \qquad\qquad\quad \mathcal{C}h\{g_j(x,\xi) \le 0\}(\alpha_j) \ge \beta_j, \quad j = 1, 2, \ldots, m \\ \qquad\qquad\quad x \in D. \end{cases} \tag{5}$$

If we want to maximize the pessimistic value to the interval random return function subject to some chance constraints, then we have the following interval random maximin chance-constrained programming model:

$$\begin{cases} \max\limits_{x} \quad \min\limits_{\bar{f}} \quad \bar{f} \\ \text{s.t.} \qquad\qquad \mathcal{C}h\{f(x,\xi) \le \bar{f}\}(\gamma) \ge \delta \\ \qquad\qquad\quad \mathcal{C}h\{g_j(x,\xi) \le 0\}(\alpha_j) \ge \beta_j, \quad j = 1, 2, \ldots, m \\ \qquad\qquad\quad x \in D. \end{cases} \tag{6}$$

3.3. Interval random simulation

In [9], we introduce two simulation algorithms to calculate the (γ, δ)-optimistic and (γ, δ)-pessimistic value to the return function $f(x, \xi)$.

According to Definition 5, the (γ, δ)-optimistic value is the maximal value $\bar{f}$ such that

$$\mathcal{C}h\{f(\xi) \ge \bar{f}\}(\alpha) \ge \beta$$

holds. It is obvious that the (α, β)-optimistic value $\bar{f}$ must be achieved at the equality case

$$Pr\{\omega \in \Omega | \mathcal{F}ea\{f(\xi(\omega)) \ge \bar{f}\} \ge \beta\} = \alpha. \tag{7}$$

We sample $\omega_1, \omega_2, \ldots, \omega_N$ from Ω according to the probability measure Pr and define

$$h(\omega_n) = \begin{cases} 1, & \text{if } \mathcal{F}ea\{f(\xi(\omega_n)) \ge \bar{f}\} \ge \beta \\ 0, & \text{otherwise} \end{cases}$$

for $n = 1, 2, \ldots, N$, which are a sequence of random variables, and $E[h(\omega_n)] = \alpha$ for all n provided that $\bar{f}$ meets (7). By the law of large numbers, we obtain $\frac{\sum_{n=1}^{N} h(\omega_n)}{N} \to \alpha$, as $N \to \infty$. Note that the sum $\sum_{n=1}^{N} h(\omega_n)$ is just the number of ω_n satisfying $\mathcal{F}ea\{f(\xi(\omega)) \ge \bar{f}\} \ge \beta$ for $n = 1, 2, \ldots, N$. Let N' be the integer part of αN. Then the value $\bar{f}$ can be taken as the N'th largest element in the sequence $\{\bar{f}_1, \bar{f}_2, \ldots, \bar{f}_N\}$ with $\bar{f}_n = \sup\{f_n | \mathcal{F}ea\{f(\xi(\omega)) \ge f_n\} \ge \beta\}$ for $n = 1, 2, \ldots, N$.

Proposition 1.

$$\bar{f}_n = m(f(\xi(\omega_n))) - w(f(\xi(\omega_n))) \times \beta.$$

Proof. $f(\xi(\omega_n))$ is a closed interval $[\underline{f}(\xi(\omega_n)), \bar{f}(\xi(\omega_n))]$, which can be obtained by interval computation. According to Eq. (2),

$$\mathcal{F}ea\{f(\xi(\omega)) \ge f_n\} = \begin{cases} 1, & f_n \le \underline{f}(\xi(\omega_n)) \\ \dfrac{m(f(\xi(\omega_n))) - f_n}{w(f(\xi(\omega_n)))}, & f_n > \underline{f}(\xi(\omega_n)), f_n < m(f(\xi(\omega_n))) \end{cases}$$

where $m(f(\xi(\omega_n))) = \frac{\underline{f}(\xi(\omega_n))+\overline{f}(\xi(\omega_n))}{2}$, $w(f(\xi(\omega_n))) = \frac{\overline{f}(\xi(\omega_n))-\underline{f}(\xi(\omega_n))}{2}$, for $n = 1, 2, \dots, N$. To meet $\mathcal{F}ea\{f(\xi(\omega)) \geq f_n\} \geq \beta$, it is obvious that $\overline{f}_n$ must be achieved at the equality case

$$\frac{m(f(\xi(\omega_n))) - f_n}{w(f(\xi(\omega_n)))} = \beta.$$

Thus,

$$\overline{f}_n = m(f(\xi(\omega_n))) - w(f(\xi(\omega_n))) \times \beta. \quad \Box$$

Similarly, we can find the (α, β)-pessimistic value to the function $f(\xi)$, which is the minimal value $\overline{f}$ such that

$$\mathcal{C}h\{f(\xi) \leq \overline{f}\}(\alpha) \geq \beta$$

holds. The difference between pessimistic value and optimistic value is that in pessimistic value

$$\mathcal{F}ea\{f(\xi(\omega)) \leq f_n\} = \begin{cases} 1, & \overline{f}(\xi(\omega_n)) \leq f_n \\ \dfrac{f_n - m(f(\xi(\omega_n)))}{w(f(\xi(\omega_n)))} \geq \beta, & \overline{f}(\xi(\omega_n)) > f_n \quad \text{and} \quad m(f(\xi(\omega_n))) < f_n. \end{cases}$$

Thus,

$$f_n \geq m(f(\xi(\omega_n))) + w(f(\xi(\omega_n))) \times \beta.$$

It is obvious that $\overline{f}_n$ must be achieved at the equality case

$$\overline{f}_n = m(f(\xi(\omega_n))) + w(f(\xi(\omega_n))) \times \beta.$$

Then the pessimistic value $\overline{f}$ can be taken as the N'th smallest element in the sequence $\{\overline{f}_1, \overline{f}_2, \dots, \overline{f}_N\}$ with $\overline{f}_n = inf\{f_n | \mathcal{F}ea\{f(\xi(\omega)) \leq f_n\} \geq \beta\}$ for $n = 1, 2, \dots, N$.

3.4. Hybrid-intelligent algorithm

A group of intelligent algorithms can be employed to find the optimal solutions of the interval random programming model (5) and (6), such as NN, GA and etc. In fact, a small change makes the hybrid NN and GA algorithm proposed in [10,11] applicable to interval random programming. A detailed description of this algorithm can be found in [10,11]. For interval random chance-constrained programming, the simulation method introduced above and new uncertain functions will substitute for the original ones in the algorithm. There are two types of uncertain function U:

$$U_1 : x \to \mathcal{C}h\{g_j(x, \xi) \leq 0, j = 1, 2, \dots, m\}(\alpha),$$
$$U_2 : x \to \sup\{\overline{f} | \mathcal{C}h\{f(x, \xi) \geq \overline{f}\}(\alpha) \geq \beta\}.$$

4. Robust portfolio selection using interval random programming

Here, We build robust portfolio model with interval random uncertainty sets based on model (1):

$$\begin{aligned} &\max_{\{x\}} \quad \min_{\{u_i \in U(u_i), \delta_{ij} \in U(\delta_{ij})\}} \sum_{i=1}^{N} u_i x_i - \lambda \sum_{i=1}^{i=N} \sum_{j=1}^{j=N} x_i x_i \delta_{ij} \\ &\text{s.t.} \quad \sum_{i=1}^{N} x_i = 1 \\ &\qquad x_i \geq 0. \end{aligned} \tag{8}$$

The program (8) is to maximize the worst-case difference of expected return and variance of portfolio.

By introducing interval random chance-constrained programming, model (8) can be transformed to the following mathematically meaningful model:

$$\begin{aligned} &\max_{\{x\}} \quad f_{\text{opt}} - \lambda g_{pes} \\ &\text{s.t.} \quad \sum_{i=1}^{N} x_i = 1 \\ &\qquad x_i \geq 0, \end{aligned} \tag{9}$$

where

$$
\begin{aligned}
f_{\text{opt}} &= \min_{\{u_i \in U(u_i)\}} \left\{ R = \sum_{i=1}^{N} u_i x_i \right\} \\
&= \max \left\{ f \,|\, \mathcal{C}h \left\{ \sum_{i=1}^{N} U(u_i) x_i \geq f \right\} (\gamma_1) \geq \phi_1 \right\}, \\
g_{pes} &= \max_{\{\delta_{ij} \in U(\delta_{ij})\}} \left\{ V = \sum_{i=1}^{i=N} \sum_{j=1}^{j=N} x_i x_i \delta_{ij} \right\} \\
&= \min \left\{ g \,|\, \mathcal{C}h \left\{ \sum_{i=1}^{i=N} \sum_{j=1}^{j=N} x_i x_i U(\delta_{ij}) \leq g \right\} (\gamma_2) \geq \phi_2 \right\}, \\
U(u_i) &= [m_i - \theta_i^1, m_i + \theta_i^2], \\
U(\delta_{ij}) &= [m_{ij} - \tau_{ij}^1, m_{ij} + \tau_{ij}^2].
\end{aligned}
$$

And $\gamma_1, \phi_1, \gamma_2, \phi_2$ are given levels. f_{opt} is the (γ_1, ϕ_1)-optimistic value to the portfolio return function $\sum_{i=1}^{N} U(u_i) x_i$. g_{pes} is the (γ_2, ϕ_2)-pessimistic value to the risk function $\sum_{i=1}^{i=N} \sum_{j=1}^{j=N} x_i x_i U(\delta_{ij})$. We have that model (9) is equivalent to

$$
\begin{aligned}
\max_{\{x\}} \quad & \max\{f\} - \lambda \min\{g\} \\
\text{s.t.} \quad & \mathcal{C}h \left\{ \sum_{i=1}^{N} U(u_i) x_i \geq f \right\} (\gamma_1) \geq \phi_1 \\
& \mathcal{C}h \left\{ \sum_{i=1}^{i=N} \sum_{j=1}^{j=N} x_i x_i U(\delta_{ij}) \leq g \right\} (\gamma_2) \geq \phi_2 \\
& \sum_{i=1}^{N} x_i = 1 \\
& x_i \geq 0.
\end{aligned}
\tag{10}
$$

5. Computational results

We test the viability of the proposed parametric approaches to mean-variance portfolio optimization using real market data. The market returns data experiments investigate whether asymmetric interval random uncertainty set is useful in real-world situations with imperfect information. We compare the performance of the portfolio model (10) taking asymmetric interval random uncertainty set to the following alternative approaches: symmetric interval random uncertainty set introduced in [9] and interval uncertainty set introduced in [12].

5.1. Experiments with real market data

We consider a portfolio of 24 small cap stocks from different industry categories of the S&P 600 index, and use 2000 daily historical returns from April, 1998 to June, 2006. Small cap stocks behave more erratically than large cap stocks, and tend to have more skewed historical return distributions. The entire data sequence is divided into time periods of length $T = 200$ days. In all there are $p = 10$ time periods. For each period p, first, we consider moving windows of $n = 10$ days and compute mean returns r'_{it} and covariance δ'_{ijt} in each such window, $i = 1, \ldots, 24$, $t = 1, \ldots, T - n + 1$ (there are $T - n + 1$ windows in each period). Then, we compute the following uncertainty sets for period $p = 1, \ldots, 10$:

(1) Asymmetric interval random uncertainty set. As defined in Section 2, the asymmetric interval random uncertainty sets take the following forms in period p:

$$
\begin{aligned}
U(u_i) &= [m_i - \theta_i^1, m_i + \theta_i^2] \\
U(\delta_{ij}) &= [m_{ij} - \tau_{ij}^1, m_{ij} + \tau_{ij}^2].
\end{aligned}
$$

Then, the uncertainty set $U(u_i)$ can be formulated as

$$
m_i = \frac{1}{T - n + 1} \sum_{t=1}^{T-n+1} r'_{it}
$$

for $i = 1, 2, \ldots, n$, in period p. Suppose θ_i^1 and θ_i^2 follow normal distribution $N(l_i^1, n_i^1)$ and $N(l_i^2, n_i^2)$, which is given by

$$l_i^1 = \frac{1}{p} \sum_{t=1}^{T-n+1} \max\{m_i - r'_{it}, 0\}$$

$$n_i^1 = \frac{1}{p} \sum_{t=1}^{T-n+1} \begin{cases} (m_i - r'_{it} - l_i^1)^2, & m_i - r'_{it} > 0 \\ 0, & m_i - r'_{it} < 0 \end{cases}$$

$$l_i^2 = \frac{1}{q} \sum_{t=1}^{T-n+1} \max\{r'_{it} - m_i, 0\}$$

$$n_i^2 = \frac{1}{q} \sum_{t=1}^{T-n+1} \begin{cases} (r'_{it} - m_i - l_i^2)^2, & r'_{it} - m_i > 0 \\ 0, & r'_{it} - m_i < 0 \end{cases}$$

where p is the number of r'_{it} where $m_i - r'_{it} > 0$. And q is the number of r'_{it} where $r'_{it} - m_i > 0$. In the same way, we can calculate the uncertainty set $U(\delta_{ij})$.

(2) Symmetric interval random uncertainty set. We assume that the uncertainty set for the expected return u_i of asset i at period p and covariance δ_{ij} of asset i and asset j at period p take the following interval random variables [9]:

$$U(u_i) = [m_i - \theta_i, m_{ip} + \theta_i]$$
$$U(\delta_{ij}) = [m_{ij} - \tau_{ij}, m_{ij} + \tau_{ij}].$$

Then, the uncertainty set $U(u_i)$ can be formulated as

$$m_i = \frac{1}{T-n+1} \sum_{t=1}^{T-n+1} r'_{it}$$

for $i = 1, 2, \ldots, n$, in period p. Suppose θ_i follows a normal distribution $N(l_i, n_i)$, which is given by

$$l_i = \frac{1}{T-n+1} \sum_{t=1}^{T-n+1} |r'_{it} - m_i|$$

$$n_i = \frac{1}{T-n+1} \sum_{t=1}^{T-n+1} (|r'_{it} - m_i| - l_i)^2.$$

In the same way, we can calculate the uncertainty set $U(\delta_{ij})$. The mean-variance model for asymmetric and symmetric will use the same model (10).

(3) Interval uncertainty set. The uncertainty set for the expected return vector u and the covariance matrix Q take the form of intervals:

$$U(u) = \{u : u^L \le u \le u^U\}$$
$$U(Q) = \{Q : Q^L \le Q \le Q^U, Q \ge 0\}$$

Here, $u^{L1}, u^{U1}, Q^{L1}, Q^{U1}$ take the 5% and 95% percentile values for mean returns r'_{it} and covariance δ'_{ijt}. $u^{L2}, u^{U2}, Q^{L2}, Q^{U2}$ take the 10% and 90% percentile values for mean returns r'_{it} and covariance δ'_{ijt}. In [12], they formulated the mean-variance model as:

$$\begin{aligned} \max_{\{x\}} \quad & u^L x - \lambda x^{\mathrm{T}} Q^U x \\ \text{s.t.} \quad & \sum_{i=1}^{N} x_i = 1 \\ & x_i \ge 0. \end{aligned} \tag{11}$$

Once all the parameters are set, the robust portfolio X_{asy}^p taking asymmetric interval random uncertainty set (resp. robust portfolio X_{sym}^p taking symmetric interval random uncertainty set, robust portfolio $X_{5\%-95\%}^p$ taking 5%–95% percentile interval, and robust portfolio $X_{10\%-90\%}^p$ taking 10%–90% percentile interval) for period p is computed by solving the robust interval random chance-constrained portfolio selection model (10) (resp. model (10), model (11)). The portfolio X_{asy}^p, X_{sym}^p, $X_{5\%-95\%}^p$ and $X_{10\%-90\%}^p$ are held constant for the period p and then rebalanced to the portfolio X_{asy}^{p+1}, X_{sym}^{p+1}, $X_{5\%-95\%}^{p+1}$ and $X_{10\%-90\%}^{p+1}$ for period $p+1$.

Let W_{asy}^p (resp. W_{sym}^p, $W_{5\%-95\%}^p$, and $W_{10\%-90\%}^p$) denote the wealth at the end of period p of an investor who has an initial wealth W^1 and employs the asymmetric measure (resp. symmetric measure, 5%–95% percentile interval, and 10%–90%

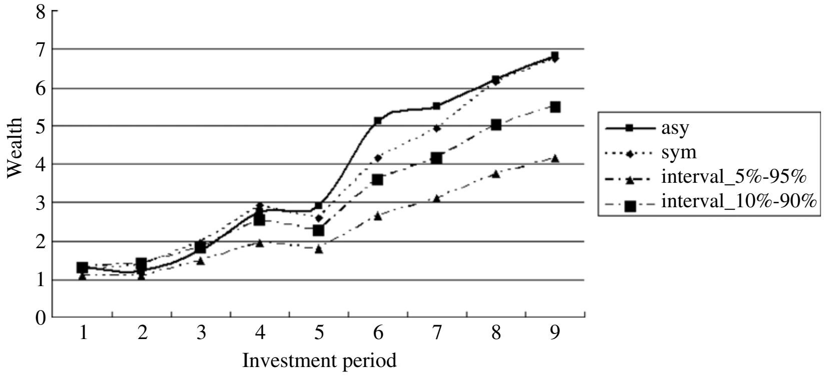

Fig. 1. The wealth resulting from the four strategies with window $n = 10$ at each investment period.

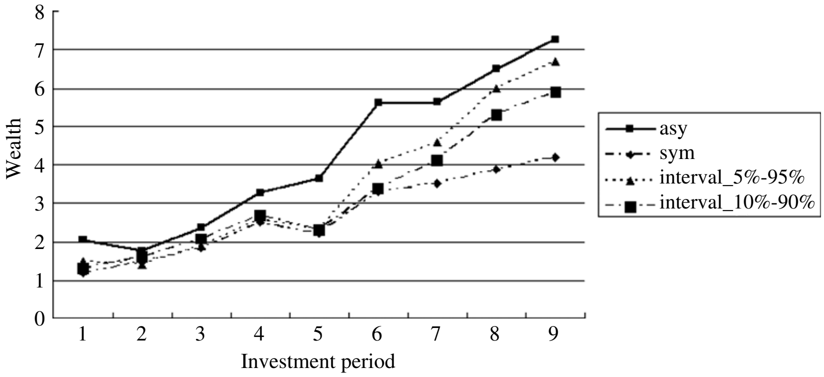

Fig. 2. The wealth resulting from the four strategies with window $n = 20$ at each investment period.

percentile interval) strategy. Here, let $W^1 = 1$. Then

$$W_{asy}^{p+1} = \left[\left(\prod_{pT<k\leq(p+1)T}(1+r_k)'\right)X_{asy}^p\right]\cdot W_{asy}^p,$$

$$W_{sym}^{p+1} = \left[\left(\prod_{pT<k\leq(p+1)T}(1+r_k)'\right)X_{sym}^p\right]\cdot W_{sym}^p,$$

$$W_{5\%-95\%}^{p+1} = \left[\left(\prod_{pT<k\leq(p+1)T}(1+r_k)'\right)X_{5\%-95\%}^p\right]\cdot W_{5\%-95\%}^p,$$

$$W_{10\%-90\%}^{p+1} = \left[\left(\prod_{pT<k\leq(p+1)T}(1+r_k)'\right)X_{10\%-90\%}^p\right]\cdot W_{10\%-90\%}^p,$$

where r_k is the vector of original daily asset returns at day k. Because these strategies require a block of data of length $T = 200$ to estimate all of parameters, the first investment period $p = 1$ starts from the time instant $T + 1$. Therefore, 10 time periods of length "$T = 200$" only have 9 investment periods.

We run the hybrid-intelligent algorithm (2000 cycles in simulation, 1000 generations in GA) to calculate the robust portfolio model (10) for 9 investment periods. Fig. 1 shows the wealth gained at the end of each investment periods for the feasibility threshold $\delta_1 = 0.9$ (resp. $\delta_2 = 0.9$) at the probability $\gamma_1 = 0.9$ (resp. $\gamma_2 = 0.9$). We also consider the moving window $n = 20$ and recompute all parameters. Fig. 2 shows the wealth with moving window $n = 20$. It is clear that the wealth generated by the asymmetric measure strategy is much better than other strategies at the end of investment period. But in Fig. 1 the wealth generated by the asymmetric measure strategy is a litter lower than other strategies at early investment periods. Therefore, it is not guaranteed that the asymmetric measure strategy always has an advantage of other three strategies.

In real world, transaction cost is another important concern for portfolio managers. When choosing investment strategy, it cannot be ignored. Here, we compare the costs of implementing the asymmetric measure strategy with those of implementing other three strategies. We calculate the transaction cost by $\|X^p - X^{p-1}\|_1$. Fig. 3 shows the ratios of the

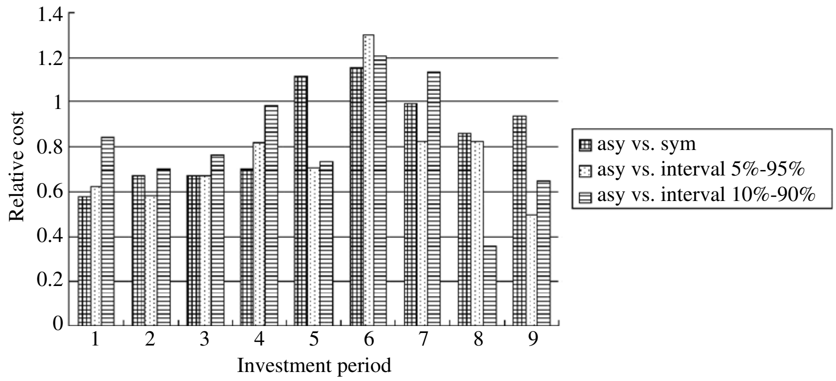

Fig. 3. Relative cost with window $n = 10$ at each investment period.

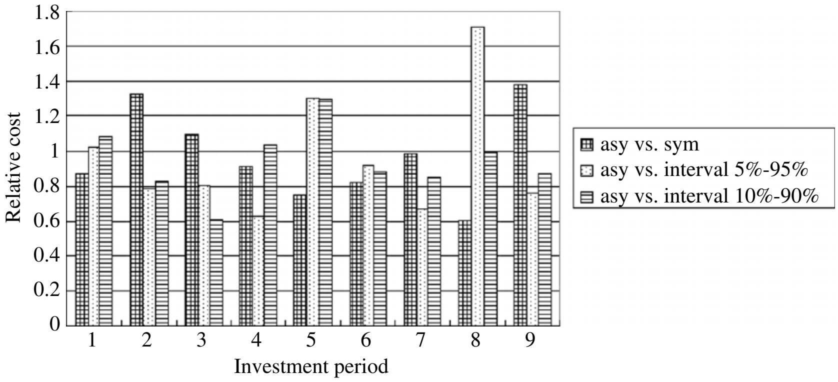

Fig. 4. Relative cost with window $n = 20$ at each investment period.

costs,

$$\|X_{asy}^{p} - X_{asy}^{p-1}\|_1 / \|X_{sym}^{p} - X_{sym}^{p-1}\|_1,$$
$$\|X_{asy}^{p} - X_{asy}^{p-1}\|_1 / \|X_{5\%-95\%}^{p} - X_{5\%-95\%}^{p-1}\|_1,$$
$$\|X_{asy}^{p} - X_{asy}^{p-1}\|_1 / \|X_{10\%-90\%}^{p} - X_{10\%-90\%}^{p-1}\|_1,$$

with moving window $n = 10$. The average costs are 0.852, 0.761, 0.819 respectively. The transaction costs of asymmetric strategy were approximately 15%, 24%, 18% less than other three strategies respectively. Fig. 4 plots the same quantity with moving window $n = 20$ and now the average costs are 0.971, 0.955, 0.938 respectively. The transaction costs of asymmetric strategy were approximately 3%, 5%, 6% less than other three strategies respectively.

6. Conclusion

Building on recent research in robust portfolio, this paper introduces a novel uncertainty set: interval random uncertainty. It can consider the variability of bounds and asymmetric measures of variability for the distribution of returns simultaneously. Robust asset allocation refers to finding an asset allocation strategy whose behavior under the worst possible realizations of the uncertain inputs is optimized. We present a robust mean-variance portfolio selection model under interval random uncertainty in the elements of mean vector and covariance matrix, and reformulate this model to a mathematically meaningful one by using our interval random chance-constrained programming. A method for generating the uncertainty set from historical data and a hybrid-intelligent algorithm for solving the robust portfolio model are discussed. The numerical experiments presented in this paper suggest that the worst-case behavior of portfolios of different asset classes can be improved significantly using the robust portfolio model under interval random uncertainty set. And the robustness is achieved at relatively high performance and low cost.

References

[1] H. Markowitz, Portfolio Selection: Efficient Diversification of Investments, Wiley, New York, 1959.
[2] A. Ben-Tal, A. Nemirovski, Robust convex optimization, Mathematics of Operations Research 23 (1998) 769–805.
[3] L. El Ghaoui, H. Lebret, Robust solutions to least-squares problems with uncertain data, SIAM Journal on Matrix Analysis and Applications 18 (1997) 1035–1064.

[4] D. Goldfarb, G. Iyengar, Robust portfolio selection problems, Mathematics of Operations Research 29 (2003) 1–38.
[5] B.V. Halldórsson, R.H. Tütüncü, An interior-point method for a class of saddle-point problems, Journal of Optimization Theory and Applications 116 (2003) 559–590.
[6] S. Ceria, R. Stubbs, Incorporating estimation errors into portfolio selection: Robust portfolio construction, Journal of Asset Management 7 (2006) 109–127.
[7] A. Ben-Tal, A. Nemirovski, Robust solutions of linear programming problems contaminated with uncertain data, Mathematical Programming 88 (2000) 411–424.
[8] F.J. Fabozzi, P.N. Kolm, D.A. Pachamanova, S.M. Focardi, Robust portfolio optimization, The Journal of Portfolio Management 33 (2007) 40–48.
[9] W. Chen, S.H. Tan, Interval random variable and interval random programming, in: review of Fuzzy Sets and Systems, 2008.
[10] B. Liu, Fuzzy random chance-constrained programming, IEEE Transactions on Fuzzy Systems 9 (2001) 713–720.
[11] B. Liu, Fuzzy random dependent-chance programming, IEEE Transactions on Fuzzy Systems 9 (2001) 721–726.
[12] R.H. Tutuncu, M. Koenig, Robust asset allocation, Annals of Operations Research 132 (2004) 157–187.